Digital Image Processing & Machine Vision

Lecture # 4

Instructed by

Dr. Abdul Rehman Abbasi

Generic Model of Machine Vision

- Scene constraints
- Image acquisition
- Preprocessing
- **Segmentation**
- Feature Extraction
- Classification and/or Interpretation
- Actuation

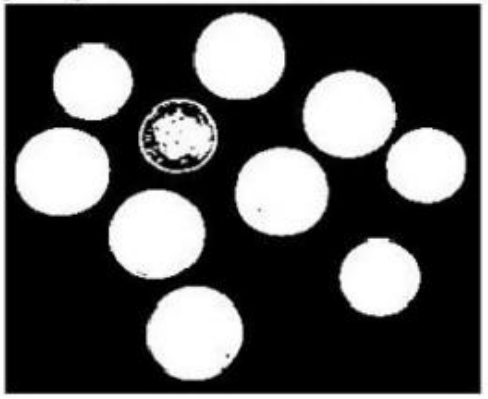
Image Segmentation-1

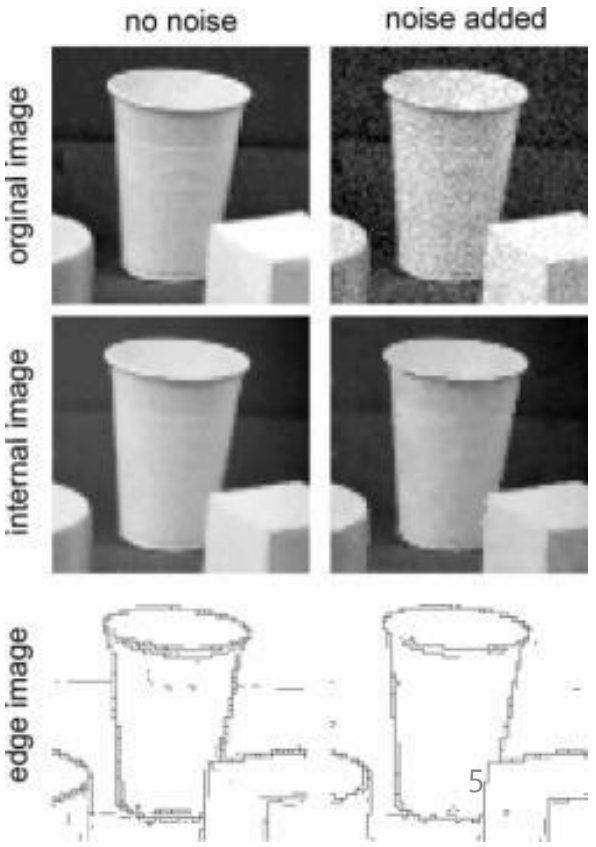
- Acquired image is broken up into meaningful regions or segments, i.e. partitioning of image.
- Segmentation is not concerned with what the image represents. Broadly two approaches are employed:
  - **Thresholding** based on some predetermined criterion (**global** thresholds the entire image into single threshold value or **local** thresholds partitions image into sub-images and determines for each of them) and
  - **Edge-based** methods (uses digital versions of standard finite operators which accentuates intensity changes, which gives rise to peak in the first derivative or a zero crossing in second derivative, which can be detected and properties such as position, sharpness, and height of peak infer the location, sharpness and contrast of intensity changes in the image).

Image Segmentation-1

- Acquired image is broken up into meaningful regions or segments, i.e. partitioning of image.
- Segmentation is not concerned with what the image represents. Broadly two approaches are employed:
  - **Thresholding** based on some predetermined criterion (**global** thresholds the entire image into single threshold value or **local** thresholds partitions image into sub-images and determines for each of them) and
  - **Edge-based** methods (uses digital versions of standard finite operators which accentuates intensity changes, which gives rise to peak in the first derivative or a zero crossing in second derivative, which can be detected and properties such as position, sharpness, and height of peak infer the location, sharpness and contrast of intensity changes in the image).

Thresholding & Edge-Based Detection





Thresholding

- Convert the images to grayscale.
- Apply a threshold to separate foreground from background.
- Use libraries like OpenCV (Python) with cv2.threshold() function.

Edge-Based Detection:

- Use techniques like Canny edge detection or Sobel filters to detect edges in the images.
- Again, OpenCV can help with cv2.Canny().

Image Segmentation-2

- The classical approach to edge-based segmentation begins with edge enhancement which makes use of digital versions of standard finite difference operators, as in the first-order gradient operators (e.g. Roberts Sobel) or in the second-order Laplacian operator.
- The difference operation accentuates intensity changes, and transforms this the image into a representation from which properties of these changes can be extracted more easily.
- A significant intensity change gives rise to a peak in the first derivative or a zero crossing in the second derivative of the smoothed intensities.
- These peaks, or zero crossings, can be detected easily, and properties such as the position, sharpness, and height of the peaks infer the location, sharpness and contrast of the intensity changes in the image.
- Edge elements can be identified from the edge-enhanced image and these can then be linked to form complete boundaries of the regions of interest

Image Segmentation-2

- The classical approach to edge-based segmentation begins with edge enhancement which makes use of digital versions of standard finite difference operators, as in the first-order gradient operators (e.g. Roberts Sobel) or in the second-order Laplacian operator.
- The difference operation accentuates intensity changes, and transforms this the image into a representation from which properties of these changes can be extracted more easily.
- A significant intensity change gives rise to a peak in the first derivative or a zero crossing in the second derivative of the smoothed intensities.
- These peaks, or zero crossings, can be detected easily, and properties such as the position, sharpness, and height of the peaks infer the location, sharpness and contrast of the intensity changes in the image.
- Edge elements can be identified from the edge-enhanced image and these can then be linked to form complete boundaries of the regions of interest

Image Segmentation-2

- The classical approach to edge-based segmentation begins with edge enhancement which makes use of digital versions of standard finite difference operators, as in the first-order gradient operators (e.g. Roberts Sobel) or in the second-order Laplacian operator.
- The difference operation accentuates intensity changes, and transforms this the image into a representation from which properties of these changes can be extracted more easily.
- A significant intensity change gives rise to a peak in the first derivative or a zero crossing in the second derivative of the smoothed intensities.
- These peaks, or zero crossings, can be detected easily, and properties such as the position, sharpness, and height of the peaks infer the location, sharpness and contrast of the intensity changes in the image.
- Edge elements can be identified from the edge-enhanced image and these can then be linked to form complete boundaries of the regions of interest

Image Segmentation-2

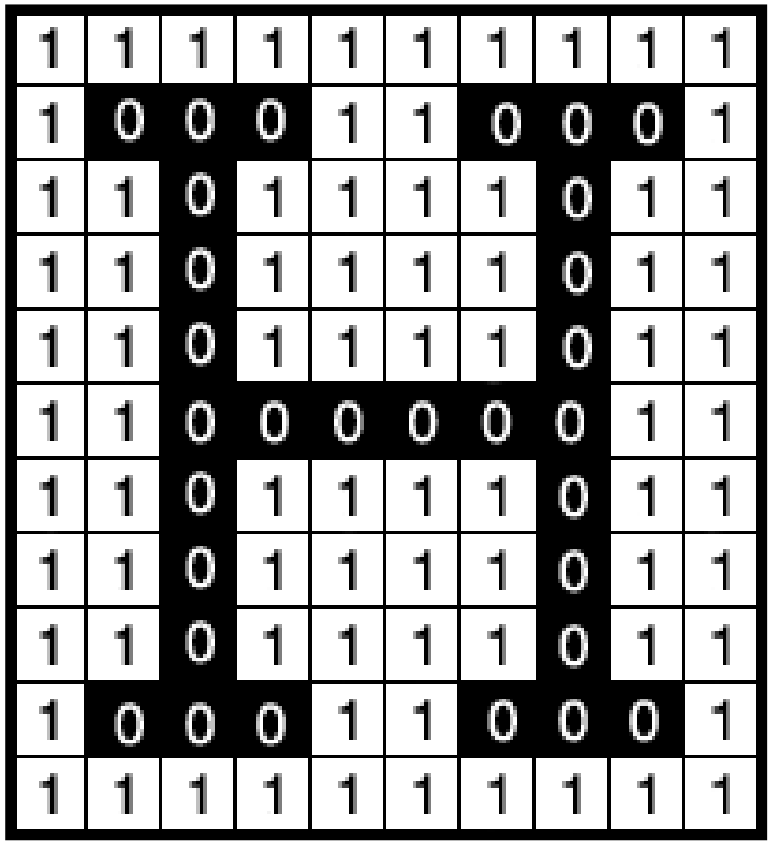
- The classical approach to edge-based segmentation begins with edge enhancement which makes use of digital versions of standard finite difference operators, as in the first-order gradient operators (e.g. Roberts Sobel) or in the second-order Laplacian operator.
- The difference operation accentuates intensity changes, and transforms this the image into a representation from which properties of these changes can be extracted more easily.
- A significant intensity change gives rise to a peak in the first derivative or a zero crossing in the second derivative of the smoothed intensities.
- These peaks, or zero crossings, can be detected easily, and properties such as the position, sharpness, and height of the peaks infer the location, sharpness and contrast of the intensity changes in the image.
- Edge elements can be identified from the edge-enhanced image and these can then be linked to form complete boundaries of the regions of interest

Image Segmentation-2

- The classical approach to edge-based segmentation begins with edge enhancement which makes use of digital versions of standard finite difference operators, as in the first-order gradient operators (e.g. Roberts Sobel) or in the second-order Laplacian operator.
- The difference operation accentuates intensity changes, and transforms this the image into a representation from which properties of these changes can be extracted more easily.
- A significant intensity change gives rise to a peak in the first derivative or a zero crossing in the second derivative of the smoothed intensities.
- These peaks, or zero crossings, can be detected easily, and properties such as the position, sharpness, and height of the peaks infer the location, sharpness and contrast of the intensity changes in the image.
- Edge elements can be identified from the edge-enhanced image and these can then be linked to form complete boundaries of the regions of interest

Image enhancement

Thresholding



Generic Model of Machine Vision

- Scene constraints
- Image acquisition
- Preprocessing
- Segmentation
- **Feature Extraction**
- Classification and/or Interpretation
- Actuation

Feature Extraction

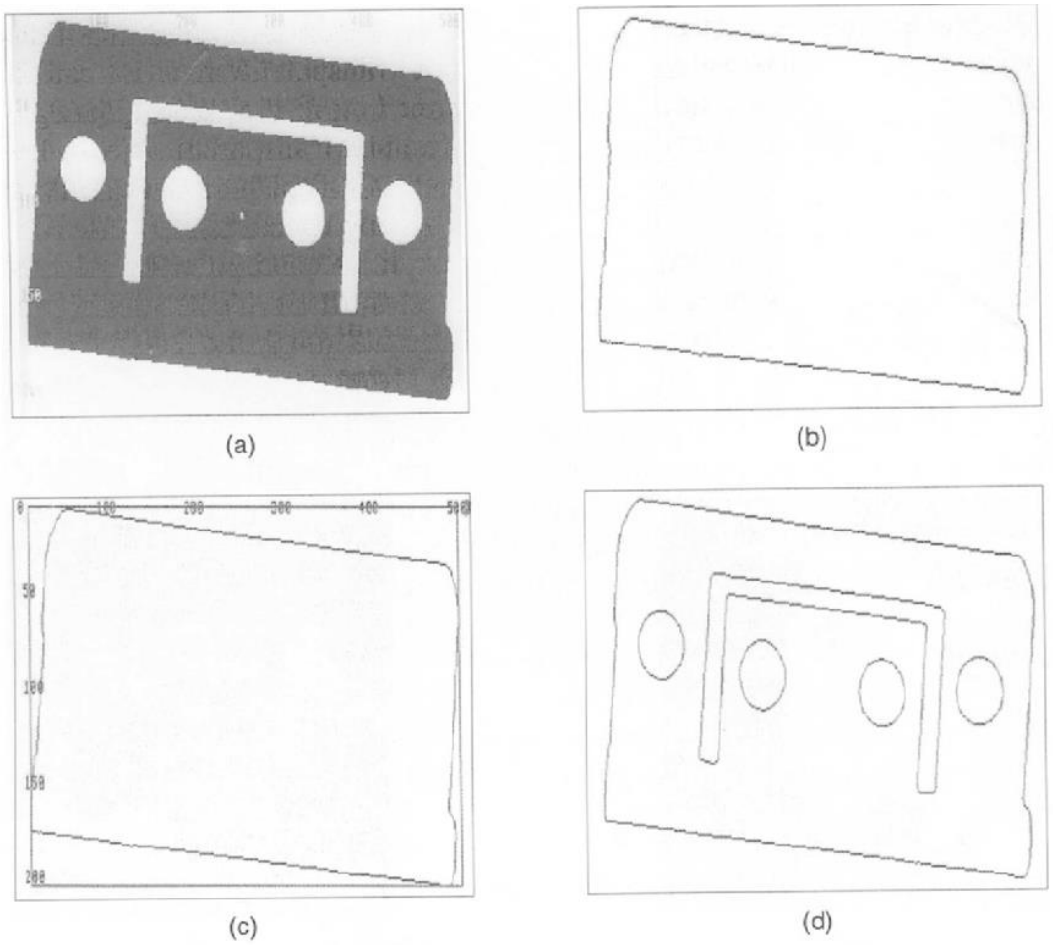


Figure 1.4 Feature extraction: (a) binary image with centroid and area data; (b) perimeter; (c) minimum enclosing rectangle; (d) number of holes identified

Feature Extraction

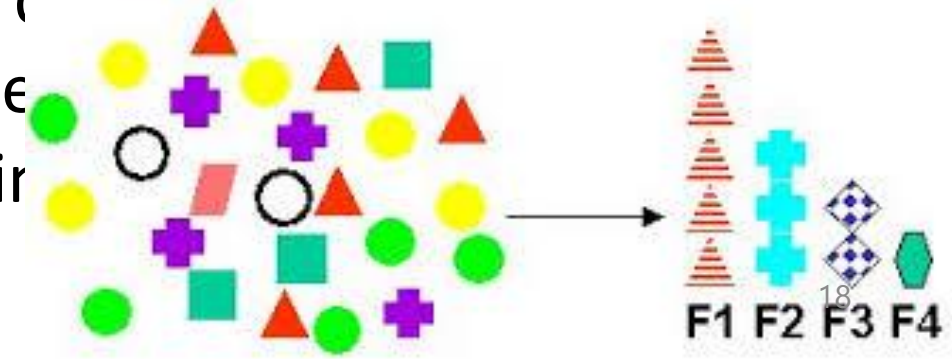
- During this phase the inherent characteristics or features of different regions within the image are identified, which are checked against predetermined standards.
- This description should be invariant to position, orientation and scale of the object.
- A number of basic parameters such as minimum enclosing rectangle, centre of area (e.g. centre may be considered as object oriented origin and series of feature descriptors can be developed), may be derived from an arbitrary shape and can be used for classification and position information

Feature Extraction

- During this phase the inherent characteristics or features of different regions within the image are identified, which are checked against predetermined standards.
- This description should be invariant to position, orientation and scale of the object.
- A number of basic parameters such as minimum enclosing rectangle, centre of area (e.g. centre may be considered as object oriented origin and series of feature descriptors can be developed), may be derived from an arbitrary shape and can be used for classification and position information

Feature Extraction

- During this phase the inherent characteristics or features of different regions within the image are identified, which are checked against predetermined standards.
- This description should be invariant to position, orientation and scale of the object.
- A number of basic parameters such as minimum enclosing rectangle, centre of area (e.g. centre may be considered as object oriented origin and series of feature descriptors can be d[illegible] may be derived from an arbitrary shape[illegible] classification and position ir[illegible]



Generic Model of Machine Vision

- Scene constraints
- Image acquisition
- Preprocessing
- Segmentation
- Feature Extraction
- **Classification and/or Interpretation**
- Actuation

Classification

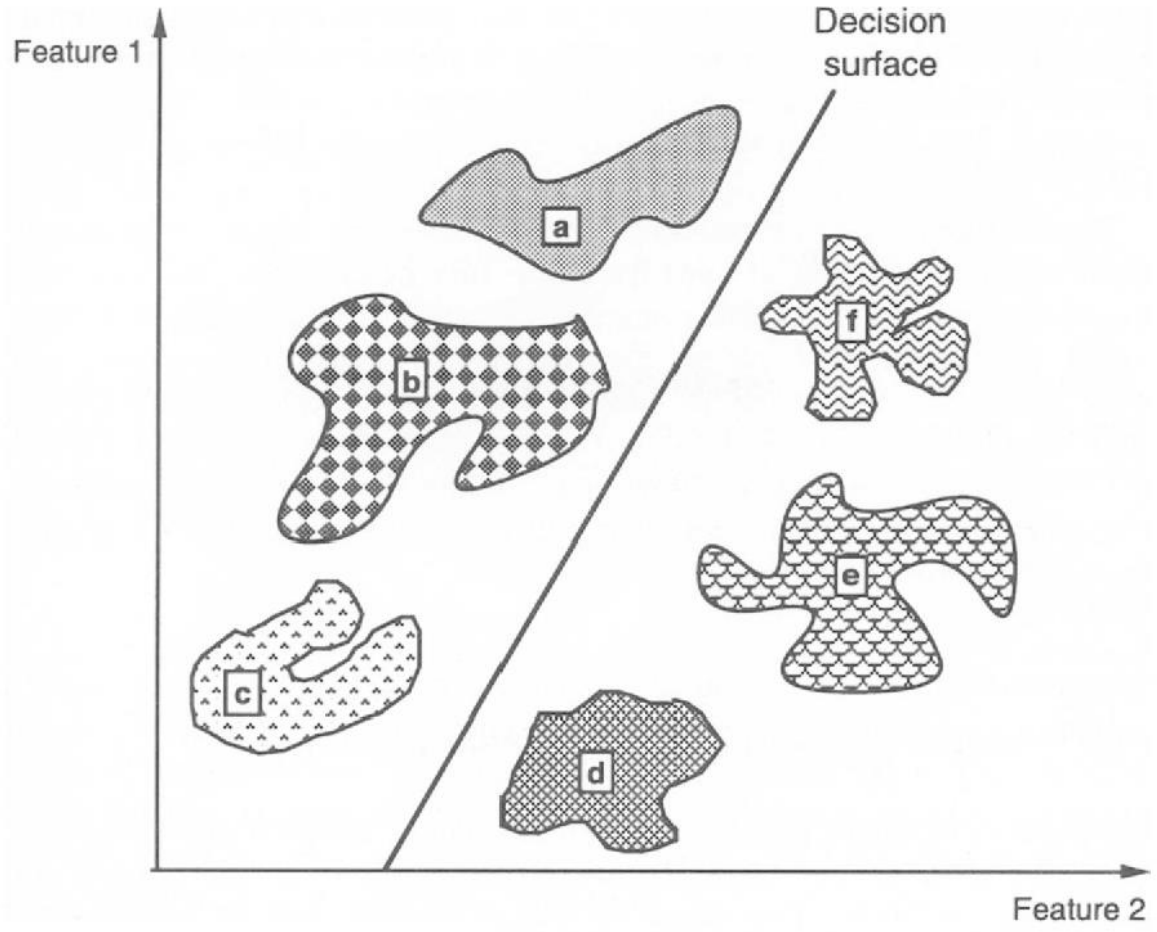


Figure 1.5 Classification via a feature space

Image Classification (Analysis)

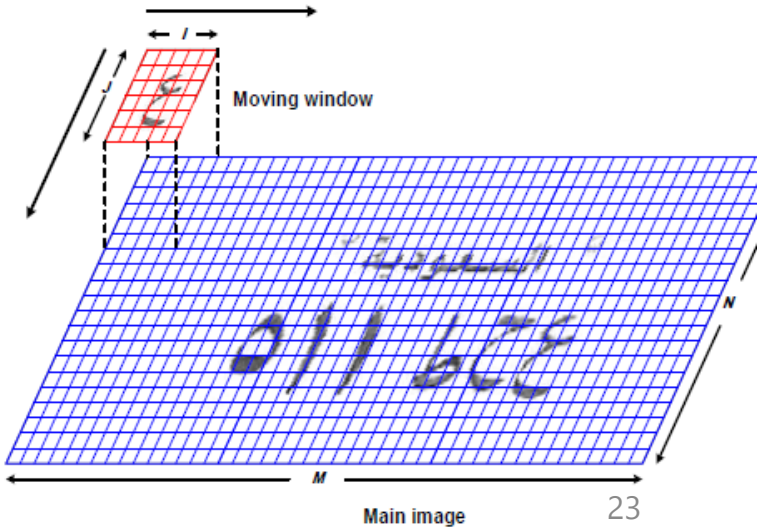
- The classification sub-system is concerned with pattern recognition or image classification. This process utilizes some or all of the extracted features to make a decision about to which category of objects the unknown object belongs.
- There are three main techniques for classification
  - Template matching
  - Statistically based approaches
  - Neural network approach

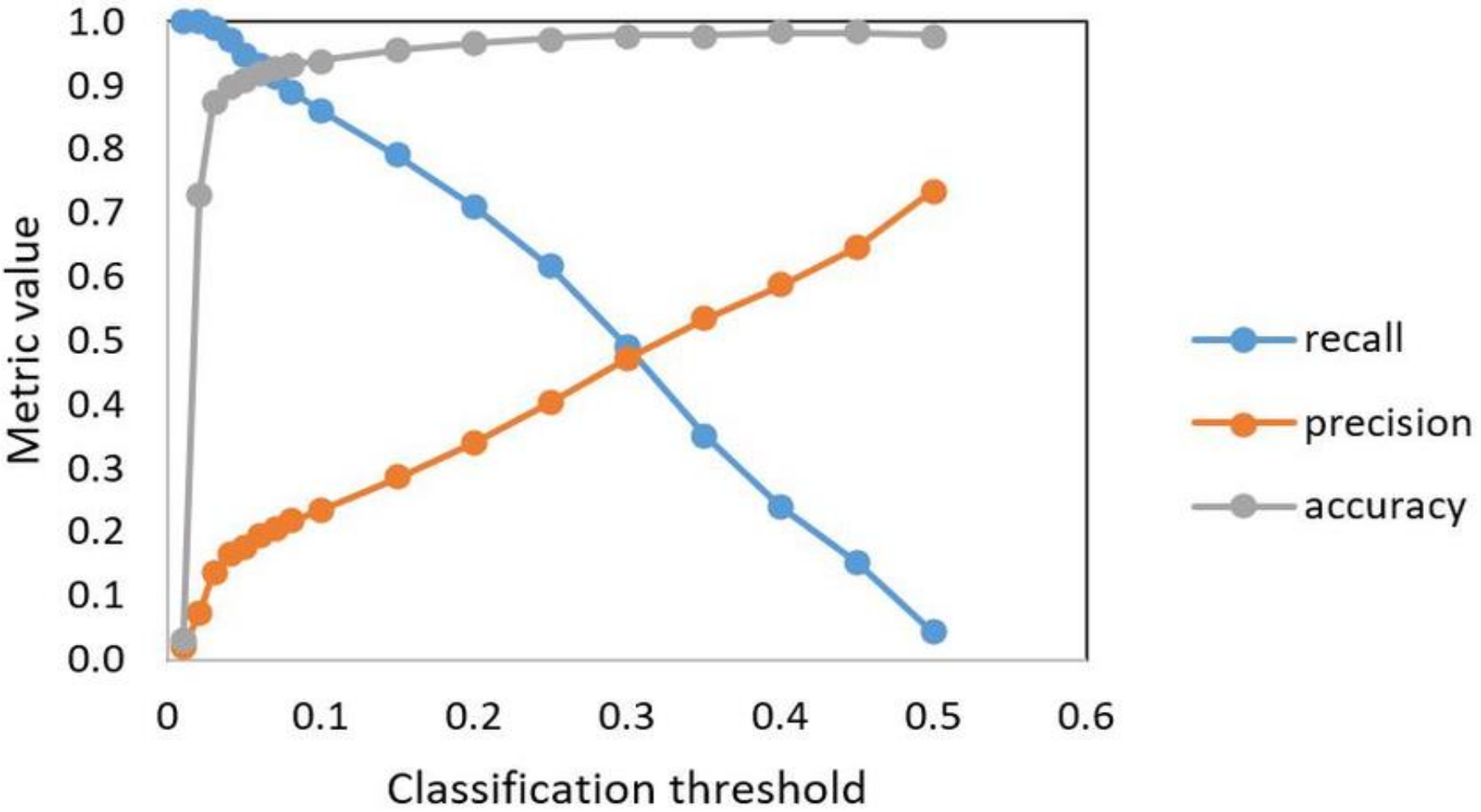
Image Classification (Analysis)

- The classification sub-system is concerned with pattern recognition or image classification. This process utilizes some or all of the extracted features to make a decision about to which category of objects the unknown object belongs.
- There are three main techniques for classification
  - Template matching
  - Statistically based approaches
  - Neural network approach

Image Classification (Analysis)

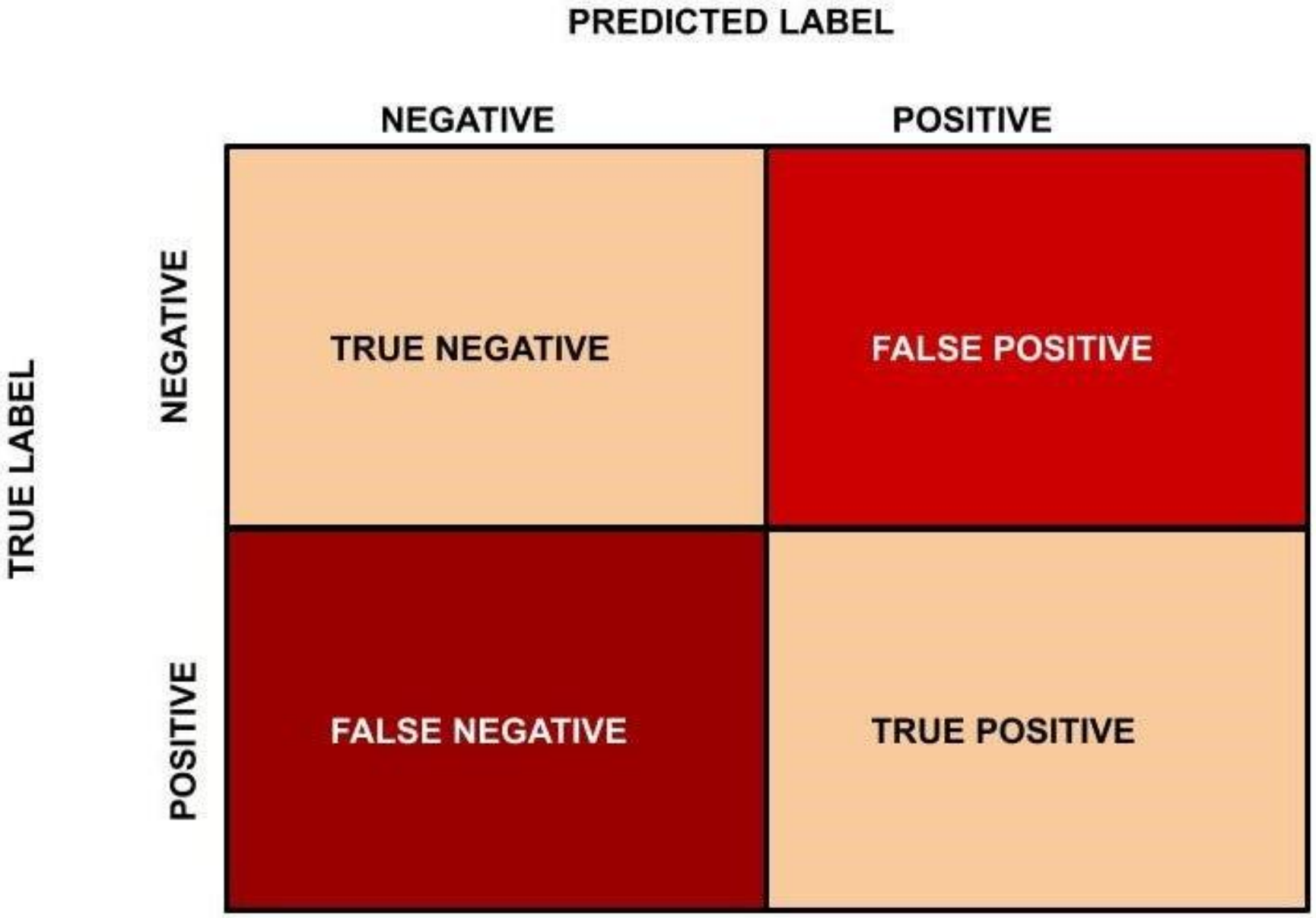
- **Template matching** is used in situations where the objects to be identified have well defined and highly 'differentiated' features, for example standard alphanumeric character fonts. In such cases an unknown character is compared with a set of templates or masks, each of which fits just one character uniquely.
- **Statistical techniques** can be selected to provide optimum classification performance for more varied industrial applications.
- If the vision task is well constrained then classification may be made via a simple tree searching algorithm where classification proceeds by making branching choices on the basis of single
complex cases, *n* features are combined to
places a candidate object within the n-dime
that the features have been properly chosen
candidate objects into well separated 'clust
consists of dividing the space with one or mo
each decision surface reliably separates two c





- **Accuracy**: Overall correctness of the model.
- **Precision**: Correctness when predicting the target class.
- **Recall**: Ability to find all objects of the target class.

Confusion Matrix

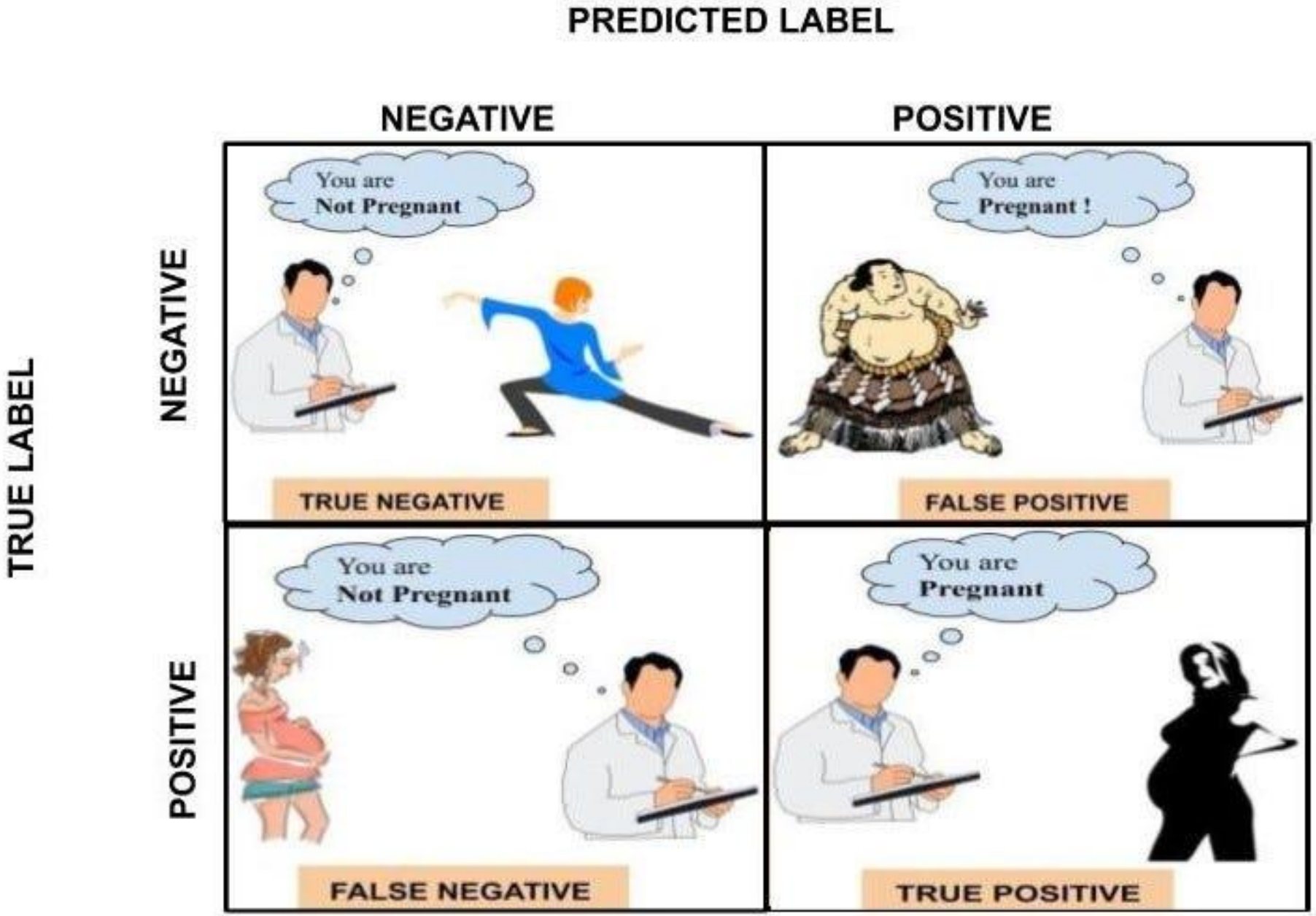

PREDICTED LABEL
NEGATIVE
POSITIVE
TRUE LABEL
NEGATIVE
POSITIVE
TRUE NEGATIVE
FALSE POSITIVE
FALSE NEGATIVE
TRUE POSITIVE

TRUE POSITIVE

TRUE NEGATIVE

FALSE POSITIVE

FALSE NEGATIVE


PREDICTED LABEL
NEGATIVE
POSITIVE
TRUE LABEL
NEGATIVE
POSITIVE
You are Not Pregnant
TRUE NEGATIVE
You are Pregnant !
FALSE POSITIVE
You are Not Pregnant
FALSE NEGATIVE
You are Pregnant
TRUE POSITIVE

$$Accuracy = \frac{TN + TP}{TN + FP + TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

Beyond Precision, Recall, Accuracy

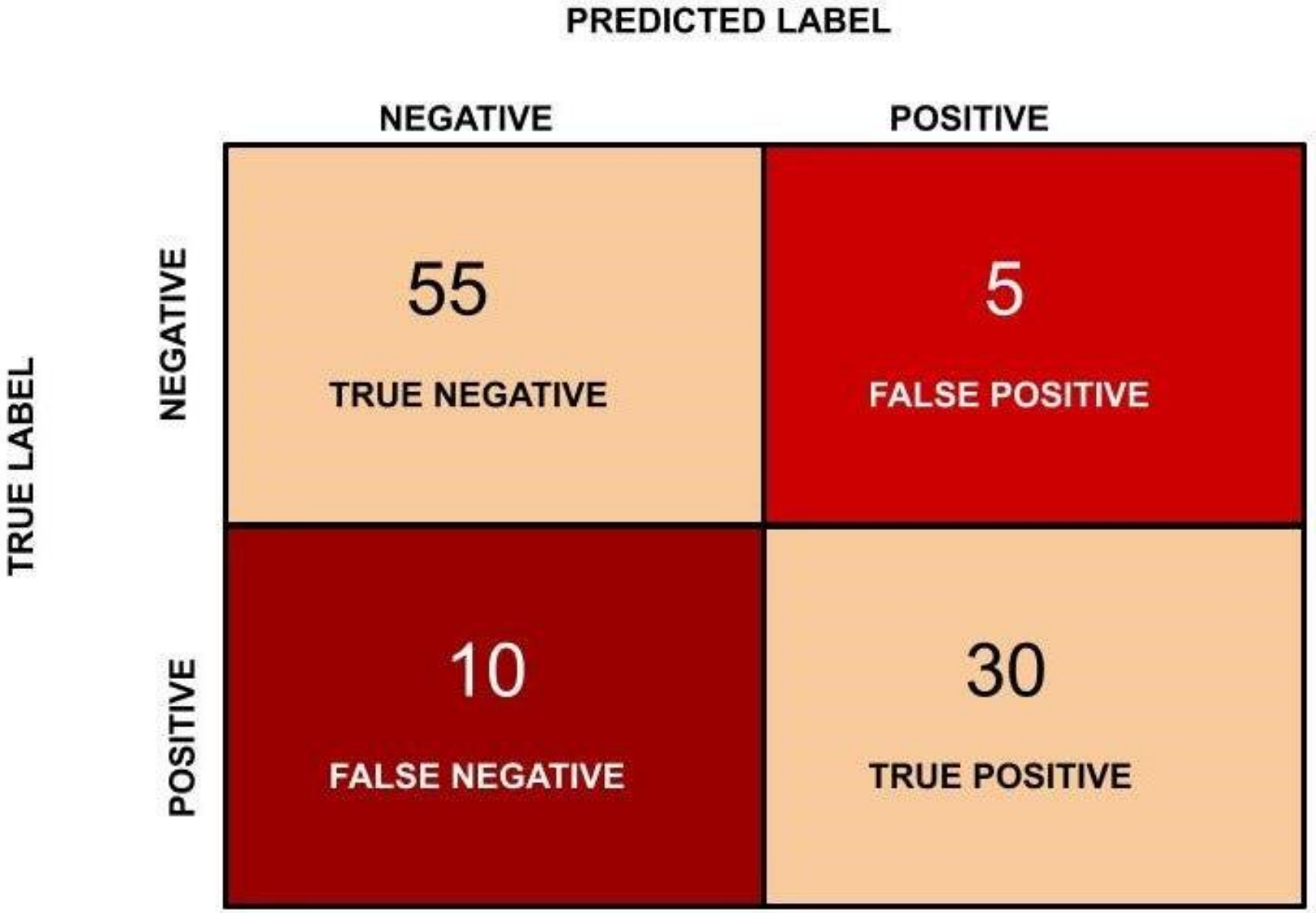
- Ideally in a good classifier, we want both ***precision*** and ***recall*** to be one which also means ***FP*** and ***FN*** are zero.
- Therefore we need a metric that takes into account both ***precision*** and ***recall***. ***F1-score*** is a metric which takes into account both ***precision*** and ***recall***.

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

- ***F1 Score*** becomes 1 only when ***precision*** and ***recall*** are both 1.
- ***F1 score*** becomes high only when both ***precision*** and ***recall*** are high.
- ***F1 score*** is the harmonic mean of ***precision*** and ***recall*** and is a better measure than ***accuracy***.

Class Quiz

- Now, we select 100 people which includes pregnant women, not pregnant women and men with fat belly. Let us assume out of this 100 people 40 are pregnant and the remaining 60 people include not pregnant women and men with fat belly. We now use a machine learning algorithm to predict the outcome. The predicted outcome (pregnancy +ve or -ve) using a machine learning algorithm is termed as the ***predicted label*** and the true outcome (in this case which we know from doctor's/expert's record) is termed as the ***true label***.
- Out of 40 pregnant women 30 pregnant women are classified correctly and the remaining 10 pregnant women are classified as not pregnant by the machine learning algorithm. On the other hand, out of 60 people in the not pregnant category, 55 are classified as not pregnant and the remaining 5 are classified as pregnant.

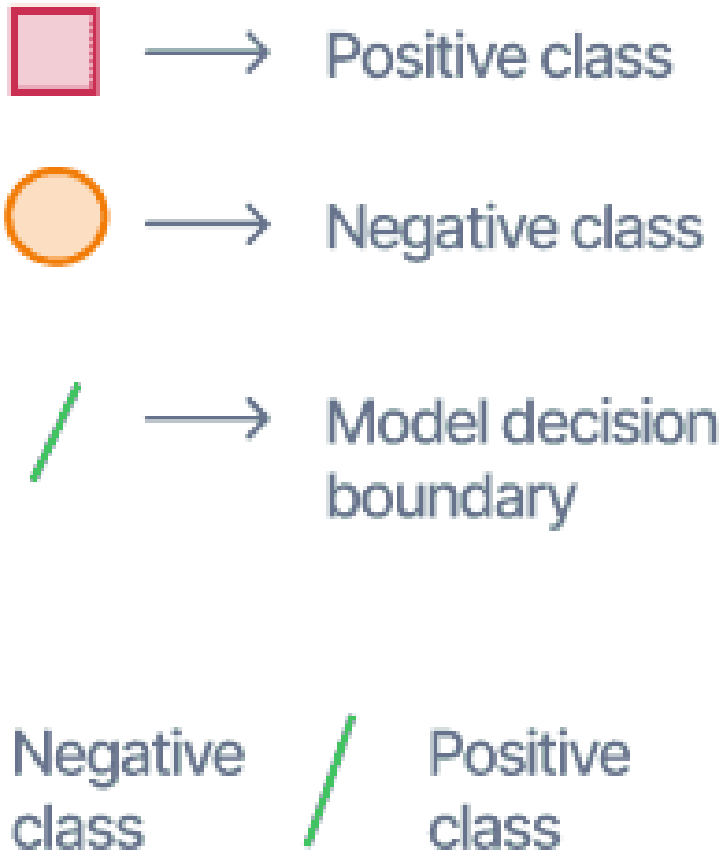

PREDICTED LABEL
NEGATIVE
POSITIVE
TRUE LABEL
NEGATIVE
POSITIVE
55
TRUE NEGATIVE
5
FALSE POSITIVE
10
FALSE NEGATIVE
30
TRUE POSITIVE

Compute

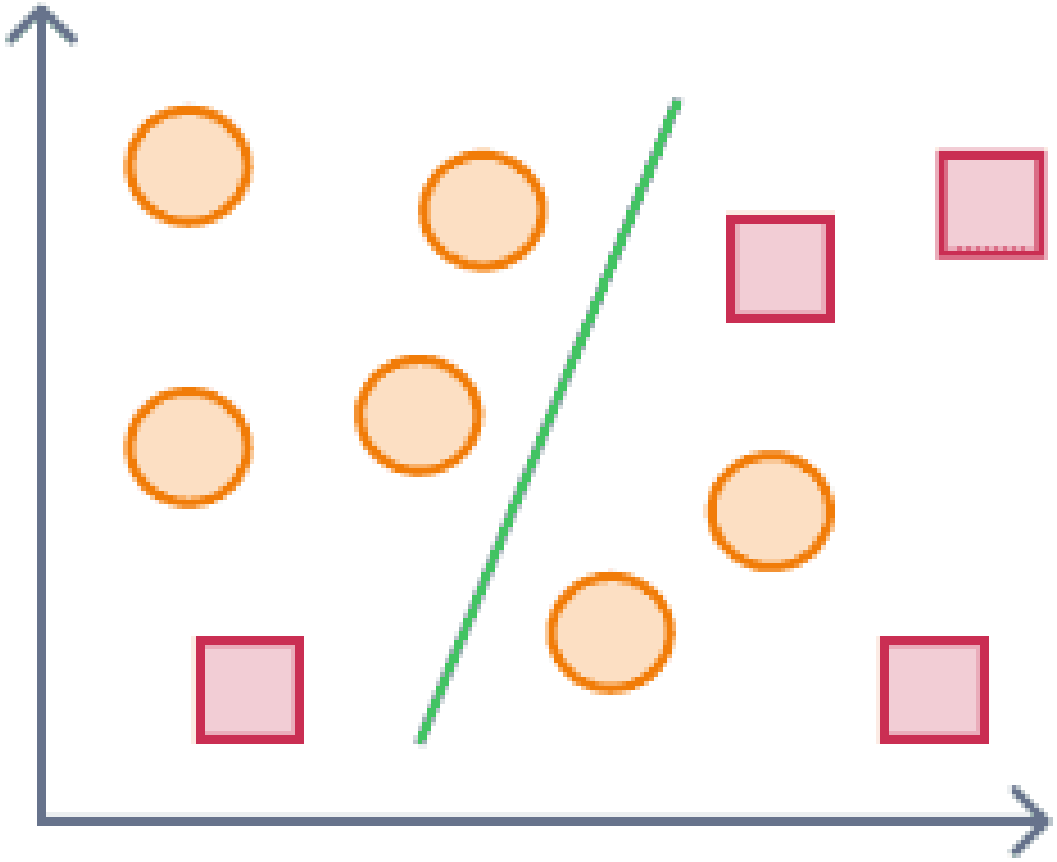
- Accuracy
- Precision
- Recall
- F1 Score

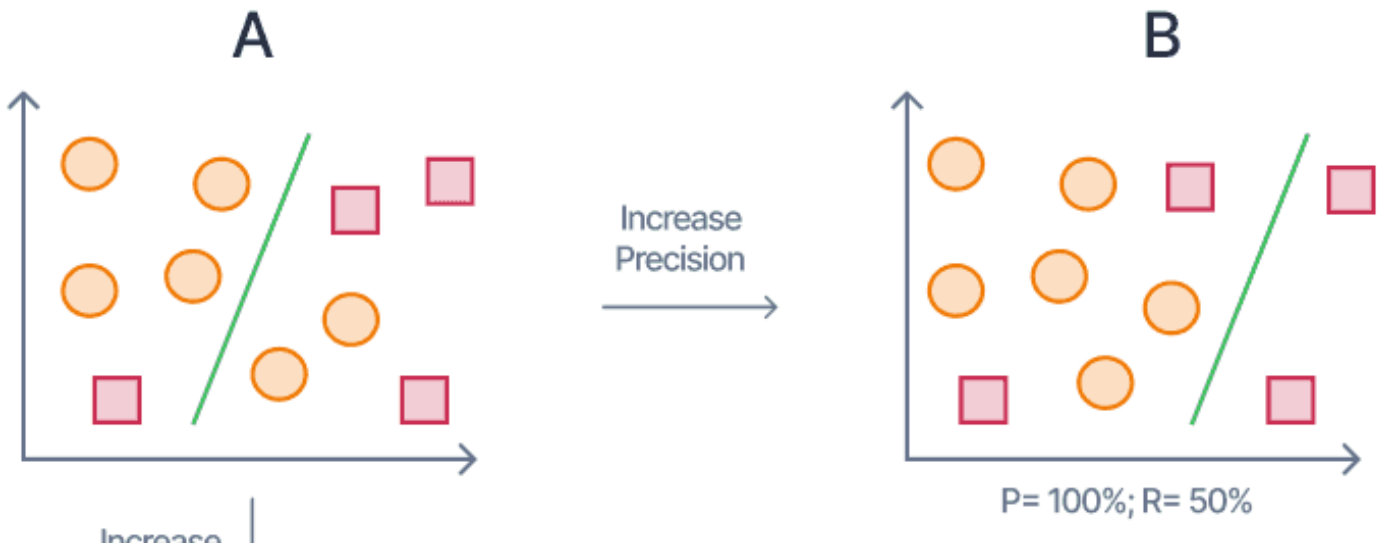
Classification

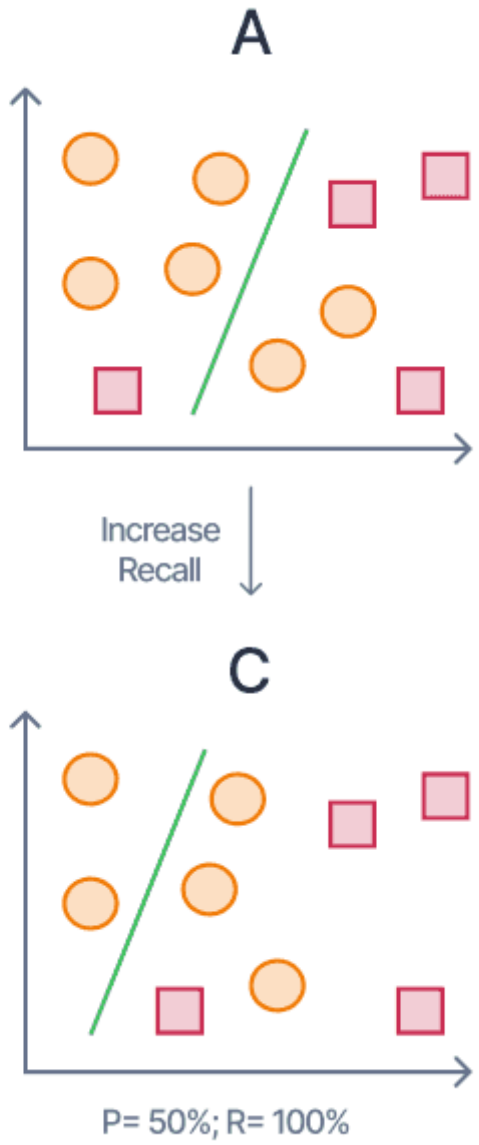
How Decision Line/Surface can affect Precision & Recall

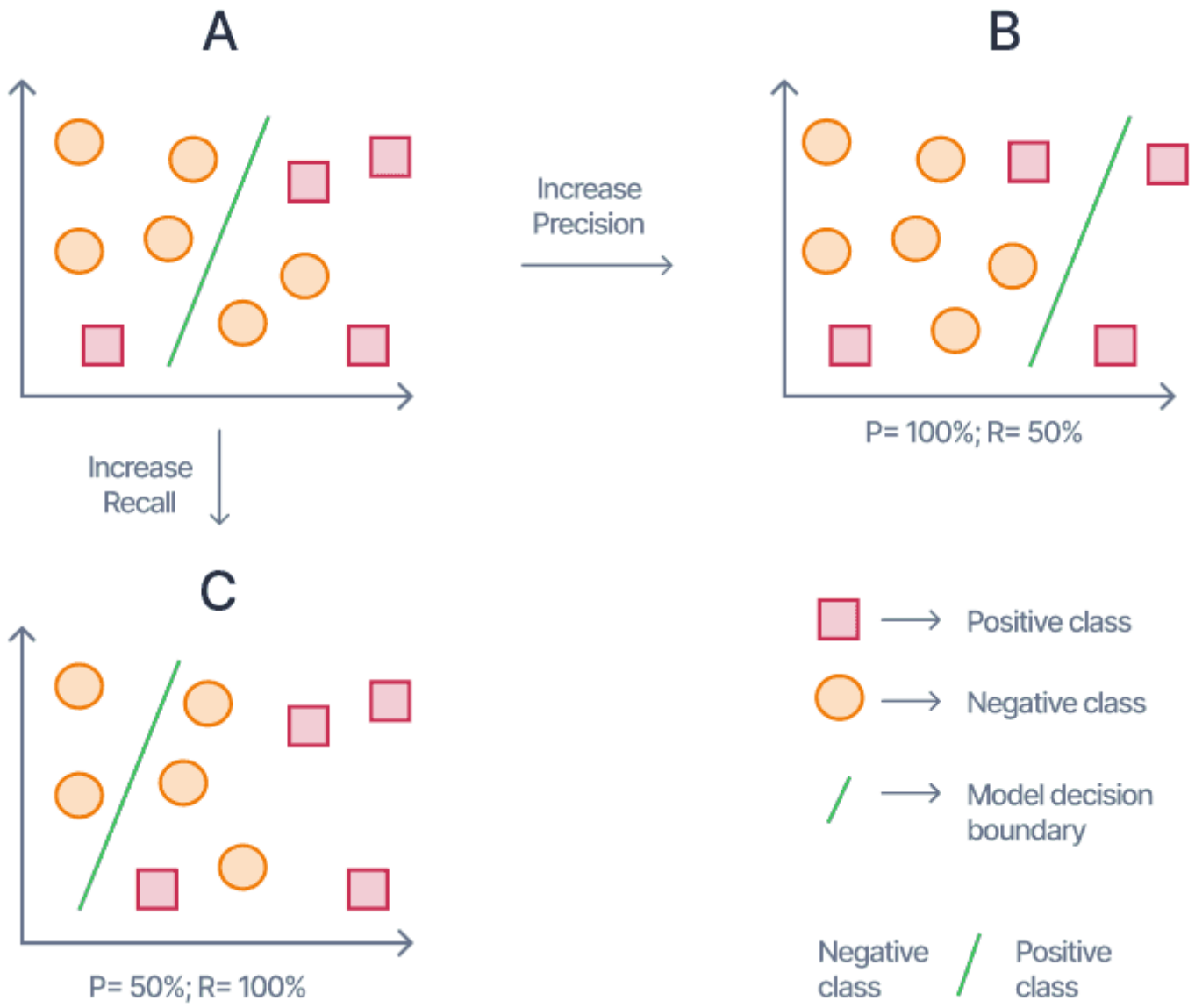

Positive class
Negative class
Model decision boundary
Negative class
Positive class

A

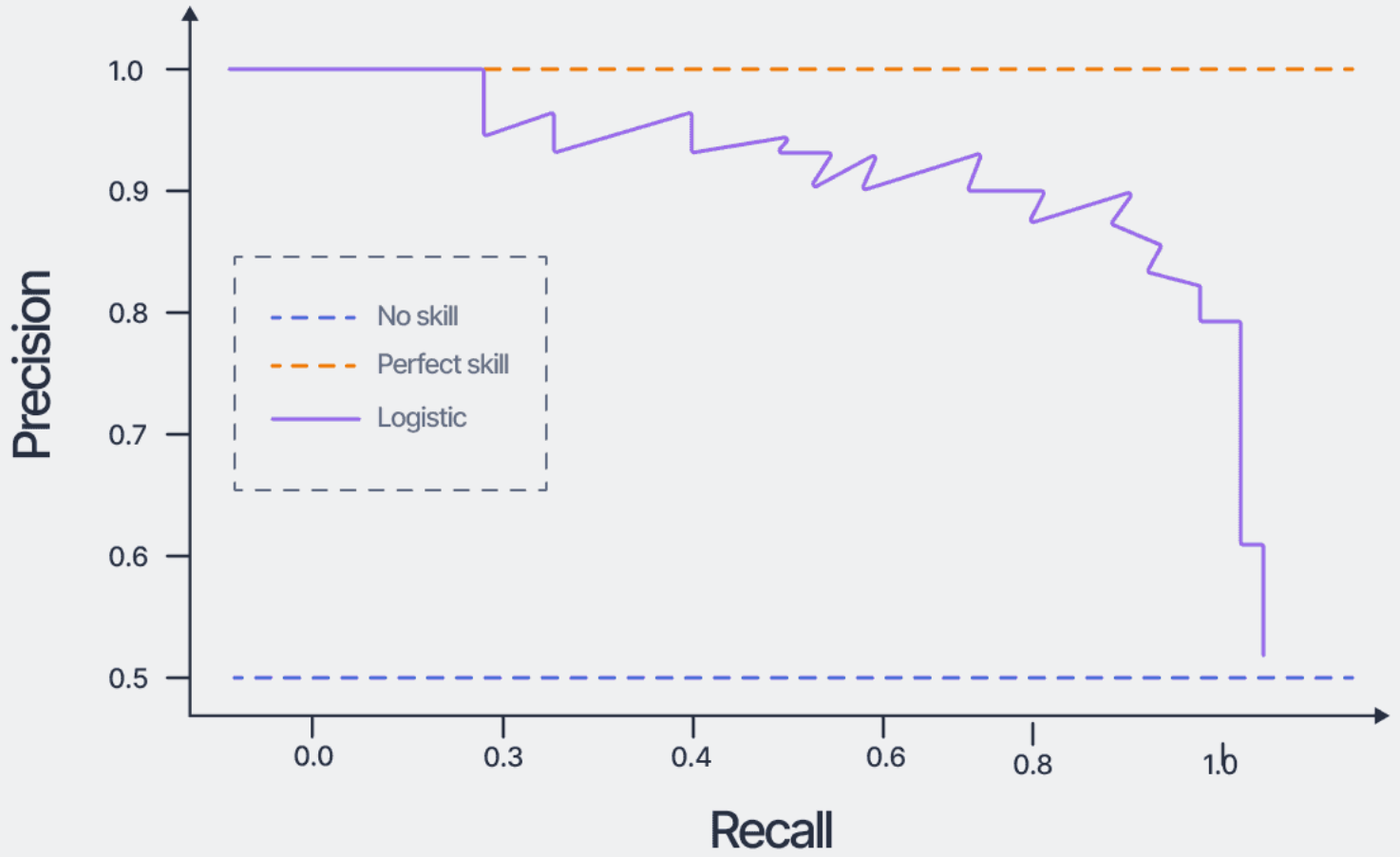



A
B
Increase
Precision
P= 100%; R= 50%


A
Increase
Recall
C
P= 50%; R= 100%


A
B
Increase
Precision
P= 100%; R= 50%
Increase
Recall
C
P= 50%; R= 100%
Positive class
Negative class
Model decision
boundary
Negative
class
Positive
class

Precision-Recall Curve



Spatial Filtering

- Some neighborhood operations work with the values of the image pixels in the neighborhood and the corresponding values of a sub-image that has the same dimensions as the neighborhood.
- The sub-image is called a filter, mask, kernel, template, or window, with the first three terms being the most prevalent terminology.
- The values in a filter sub-image are referred to as coefficients, rather than pixels.
- The term spatial filtering is used to differentiate from the more traditional frequency domain filtering

Spatial Filtering

- For linear spatial filtering, the response is given by a sum of products of the filter coefficients and the corresponding image pixels in the area spanned by the filter mask

$$R = w(-1,-1)f(x-1,y-1) + w(-1,0)f(x-1,y) + \cdots + w(0,0)f(x,y) + \cdots + w(1,0)f(x+1,y) + w(1,1)f(x+1,y+1),$$

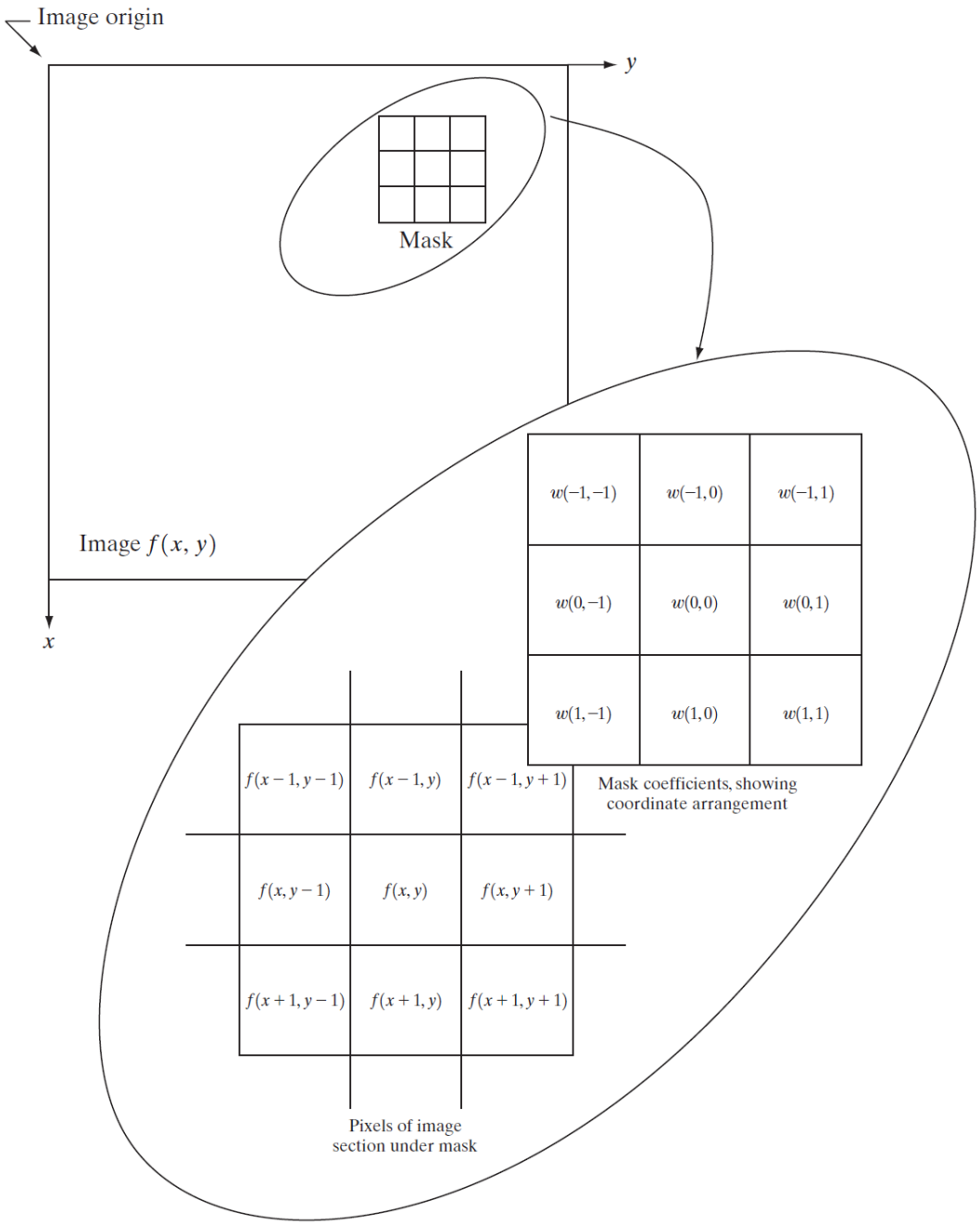
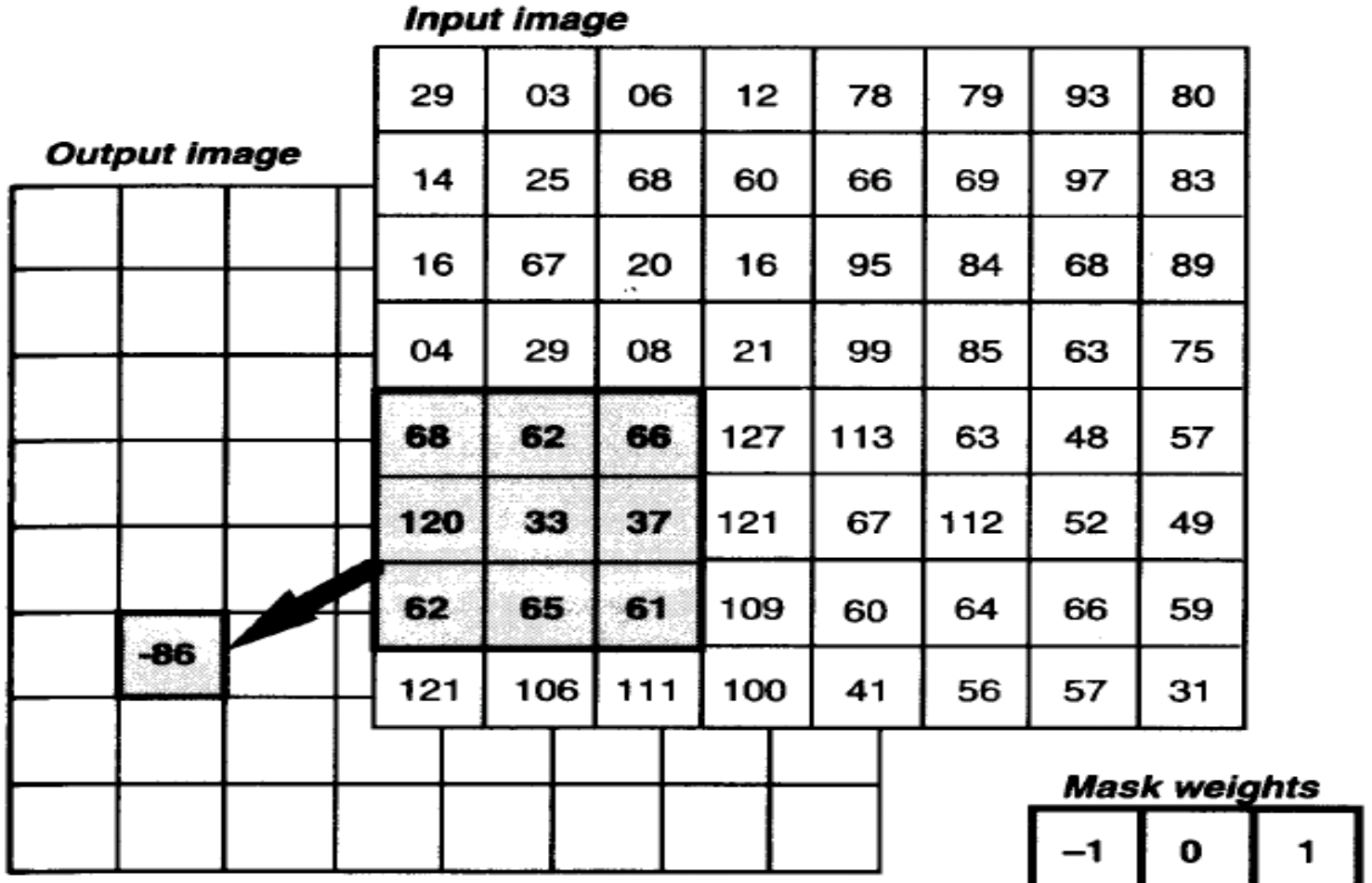


FIGURE 3.32 The mechanics of spatial filtering. The magnified drawing shows a 3×3 mask and the image section directly under it; the image section is shown displaced out from under the mask for ease of readability.



Mask weights

–1	0	1
–1	0	1
–1	0	1

Output value is -86, because:

(–1x68) + (0x62) + (1x66) + (–1x120) + (0x33)
+ (1x37) + (–1x62) + (0x65) + (1x61)
= –86

Spatial Filtering

w(-1,-1)	w(-1,0)	w(-1,1)
w(0,-1)	w(0,0)	w(0,1)
w(1,-1)	w(1,0)	w(1,1)

Mask coefficients showing coordinate arrangement

Convolution

$$g(x,y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} \omega(s,t) f(x-s, y-t)$$

$$g = \omega * f$$

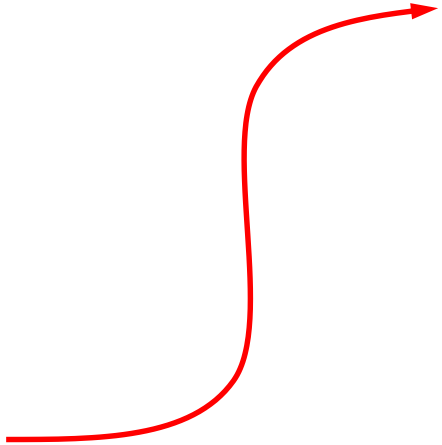
Convolution

Convolution kernel, ω

1	-1	-1
1	2	-1
1	1	1

Rotate 180°

1	1	1
-1	2	1
-1	-1	1



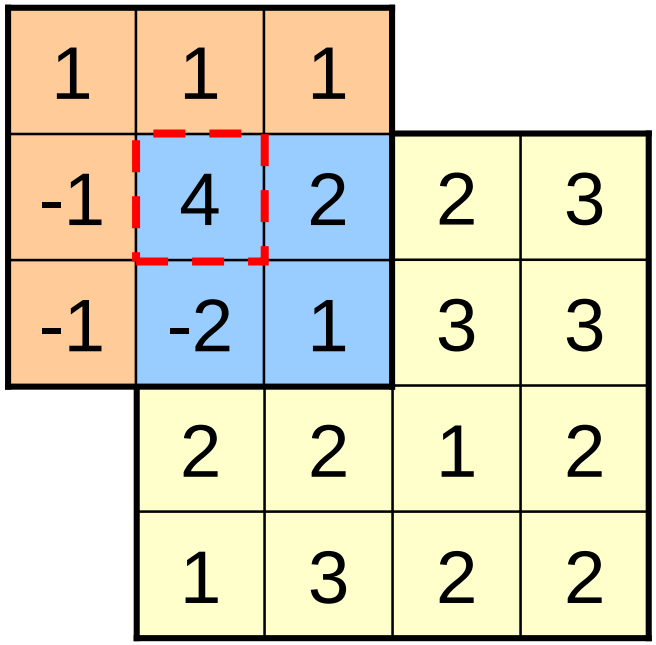
Input Image, f

2	2	2	3
2	1	3	3
2	2	1	2
1	3	2	2

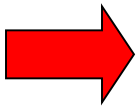
Convolution

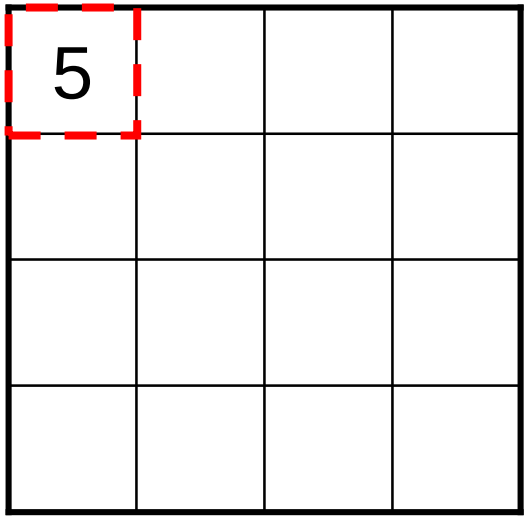
1	1	1
-1	2	1
-1	-1	1

2	2	2	3
2	1	3	3
2	2	1	2
1	3	2	2



Input Image, *f*





Output Image, *g*

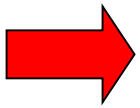
Convolution

1	1	1
-1	2	1
-1	-1	1

2	2	2	3
2	1	3	3
2	2	1	2
1	3	2	2

1	1	1	
-2	4	2	3
-2	-1	3	3
2	2	1	2
1	3	2	2

Input Image, *f*



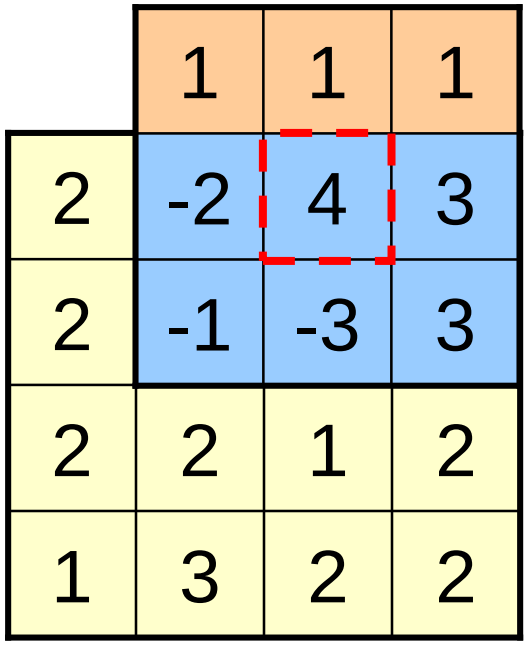
5	4		

Output Image, *g*

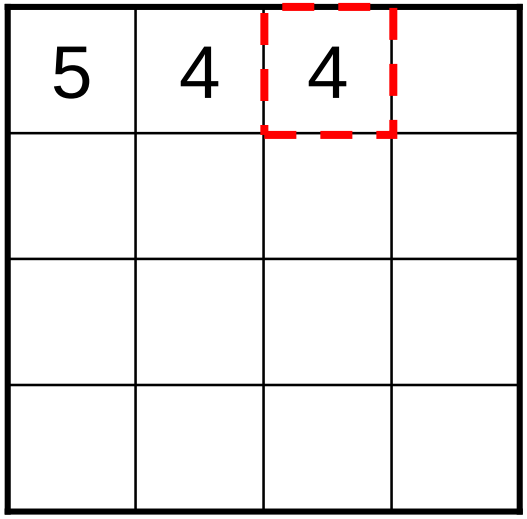
Convolution

1	1	1
-1	2	1
-1	-1	1

2	2	2	3
2	1	3	3
2	2	1	2
1	3	2	2



Input Image, *f*

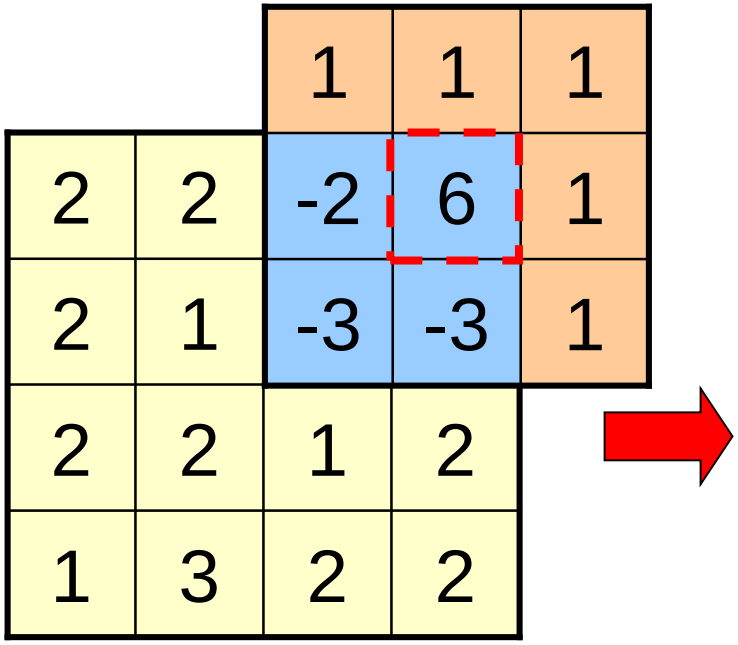


Output Image, *g*

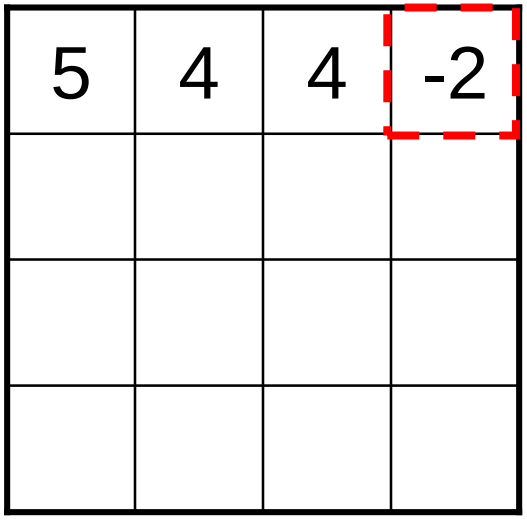
Convolution

1	1	1
-1	2	1
-1	-1	1

2	2	2	3
2	1	3	3
2	2	1	2
1	3	2	2



Input Image, *f*



Output Image, *g*

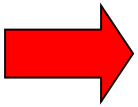
Convolution

1	1	1
-1	2	1
-1	-1	1

2	2	2	3
2	1	3	3
2	2	1	2
1	3	2	2

1	2	2	2	3
-1	4	1	3	3
-1	-2	2	1	2
	1	3	2	2

Input Image, *f*



5	4	4	-2
9			

Output Image, *g*

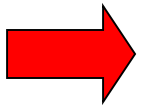
Convolution

1	1	1
-1	2	1
-1	-1	1

2	2	2	3
2	1	3	3
2	2	1	2
1	3	2	2

2	2	2	3
-2	2	3	3
-2	-2	1	2
1	3	2	2

Input Image, *f*



5	4	4	-2
9	6		

Output Image, *g*

Convolution

5	4	4	-2
9	6	14	5
11	7	6	5
9	12	8	5

Final output Image, *g*

Correlation

$$g(x,y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} \omega(s,t) f(x+s, y+t)$$

$$g = \omega \circ f$$